Personalized Privacy Assistant To Capture, Communicate And Enforce Privacy preferences and Raise awareness on privacy

G.D.S.T.Rathnasekara[1] and PPNV Kumara[2]

[1,2] *Faculty of Computing, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka*

[1] shanakagdtr@gmail.com, [2] nandana@kdu.ac.lk

Abstract— *With the rapid expansion and evolvement of the field of computing number of connected devices and number of people getting access to internet are rising in an exorbitant rate day by day. All these IOT devices and people are generating huge amounts of data. Mobile phones, big data, internet of things (IOT) are making it totally impractical for people to be aware of the means in which their data can potentially be collected, processed and used. Because of this overwhelming number of privacy related threats and issues are rising day by day. Users are unable to adequately manage their privacy settings. The aim of this study is to investigate the different privacy related issues, how those issues arise and a solution to alleviate these rising issues. People are in need of an intuitive, effective and easy to manage solution to alleviate these issues with their own minimum intervention. According to literatures, most concerning privacy issues are related with mobile phones, internet of things devices and online tracking. Finally, the study conclude that Personalized Privacy Assistant is the optimum solution for all these concerns. This paper had summarized the most important core functionalities that the privacy assistant must include for the android platform in order to cater the ever-expanding number of privacy related issues and concerns.*

Keywords— **Privacy Preferences, Personalized Privacy Assistant, Raising awareness on privacy, Privacy Concerns, Android**

I. INTRODUCTION

Due to the rapid evolvement and expansion of technology, more and more people gain access to the internet by means of a computing device. Out of the 7.6 billion people, 4.2 billion have some sort of access to the internet. All these users are leaving a digital footprint in each and every website, application and service they use and interact with. Apart from people the Internet of Things are also generating huge amounts of data. So developers are using different techniques in order to make use of these digital footprints left by the users which are then processed in order to identify the users behaviours, preferences, personal overviews, personal insights etc. The data generated by the IOT devices are also used by the developers to harness some important metrics.

Mobile phone adoption, IOT devices and big data are the main reasons for the rising number of privacy related issues in the present world. Online trackers are used in order to make targeted advertisement for the specific users, in order to customize websites and search results, tailored recommendations and better understand the userbase through analytics. So this has been a widespread practice for an increasingly personalized user experience. But people are lacking the awareness about the background processes that deliver these services.

There is an ever expanding eco system build around mobile phones. Smartphone had revolutionized the way people access internet and the computing industry is having an exponential growth due to the adoption of mobile phones. Smart phone applications are used to enhance the functionalities and make the peoples day to day life easier. These applications consist of a privacy policy which are intended to describe the smartphone app's data collection and usage practices. However not all apps have privacy policies. Its hard to evaluate the apps' privacy practices for users, regulators and privacy organizations without a proper privacy policy. The mobile applications request certain permissions. The permissions are mainly requested in order to deliver a core functionality of the application orelse for the purpose of analytics or to serve personalized information with the advertising networks. There are also a growing number of API's involve in accessing sensitive functional data of users. There are over 130 permissions in the android eco system. So its really impractical for the users to review and adjust all these permission settings on their own.

There is a rising need of better understanding of how the people feel about the privacy implications of IoT and the circumstances where they prefer to be notified about the data collection. In the IoT field there is a need for transparency, control and new methods to ensure that individual privacy requirements are met.

So the above discussed privacy issues and concerns had highlighted the need of this sort of a research around privacy practices. There is an urgent need for a proper solution to overcome these issues without a major intervention from the users end. There had been a substantial amount of research efforts devoted into finding an adaptive and effective solution for this. So there is a rising need of a more scalable solution that can empower users to regain appropriate control of their data.

II. LITERATURE REVIEW

A. Online Tracking

Online tracking has become a widespread practice that is used in most of the fields in computing(Melicher et al., 2016). Particularly the internet requires the users to disclose their personal information online for various reasons. The main intention behind online tracking is to have a more personalized user experience and more effective methods of advertising. Due to this users have significant privacy concerns and the policy initiatives have trouble in establishing a proper guideline for this. Most studies have suggested that users preferences in this regard are so complex and diverse making it really difficult to capture and categorize(Joinson et al., 2010). The research community had achieved a certain general understanding in this regard but a precise understanding of these factors are necessary in order to develop an effective technical solution. There must be a proper examination of actual web browsing situation of user in order to get the inner workings of these.

B. Mobile Phones

Smart phone adoption had revolutionized the computing industry in the past decade. More and more people got hold of a computing device through the wide adoption and massive cost reduction of this portable computing device. To extend the functionality of the smart phone, applications are developed by various developers. The growth of the number of mobile apps had also fuelled by the increasing number of APIs made available to developers(Liu et al., 2014). These APIs can access the sensitive information of the users such as current location, contact list, call logs likewise. These are the reasons behind the rise of the security and privacy risks that had risen in the present. According to the recent studies it states that user's willingness to grant a given permission to a given mobile app is strongly influenced by the purpose associated with the certain permission(Balebako et al., 2014). As an example, granting of permission to access location data by a certain application will depend on whether it is needed by one of the core functionalities in the application or else for the purpose of analytics or serving advertisements. There is a restriction in the android applications that the sensitive resources can only be accessed if the corresponding permissions are declared in their manifest files and obtain authorization from users to use them at the time of installation.

When considering the plethora of permissions in mobile applications and users have a totally diverse set of privacy preferences.(Lin et al., n.d.) These preferences can be clustered together into few profiles which will fit with the entire population. This will save a lot of tedious work for the users, so that they don't need to review each and every permission.

Privacy nudges are great method to keep the users informed about the ongoing workings in the background. By using the nudges with the appropriate details is so much important for the users. So this will effectively increase the user's privacy awareness and motivate them to review and adjust their permissions.

C. Internet Of Things Devices

When considering the Internet of Things devices, they are consisted of physical devices from sensors that people voluntarily wear or carry with them to network connected thermostats and the street lights that counts the number of people that pass by. These devices are bringing new services, increase convenience for the humanity, improve efficiency and in return they are also bringing serious privacy and security risks. (Emami-Naeini et al., n.d.) So the best way is to give the insight to the users about the process of data collection and usage of them. Some people are fine with cameras or CCTV recordings of outdoors in public spaces but they express contempt when installing video surveillance inside the home(Atzori et al., 2010). Its totally an individual's privacy preference. Most of the people may feel comfortable with location being tracked for the purpose of traffic prediction but may consent it if their location data is retained and used in an anonymized form. Aggregated data from the IOT devices can provide a wealth of knowledge for important aspects like disaster management, customer sentiment analysis, smart cities and bio surveillance.

Based on (Perera et al., n.d.) researches and IT professionals are paying more attention towards the IoT technologies, business models associated with them and potential regulatory efforts to ensure that a more secure and privacy preserved IoT data management techniques are developed. The sheer number of IoT devices that will be installed in the future is the most crucial factor that will led to the massive rise in privacy threats. Since there might be more and more devices if there is no proper regulation in the future there will be a huge threat in connection with privacy issues due to IoT devices

III. DISCUSSION

According to the above literatures, most concerning privacy issues are related with the internet of things devices, smart phones and online tracking. So these privacy issues must be minimized by developing an effective and intuitive personalized privacy assistant.

A. Privacy in the Internet of things

So as discussed above the internet of things are the network of network of massive number of objects, sensors or devices. By 2020 there might be 50 to 100 billion devices connected to the internet. So these will produce massive amount of big data for analysis and knowledge extraction.

In the future in IOT era there may be two models on which the data will be handle(Leon et al., 2013)

1. Some consumers will be willingly to pay to use services to ensure that their privacy will be protected
2. Rest will provide the data under some limitations and conditions

Some common research questions that the researches came across were categorized into 5 main categories.(Liu et al., 2014)

- User Consent Acquisition – How the privacy policies and terms related to IoT are presented to the users.
- Control, Customization and Freedom of choice – How the users are given the control and management of their data
- Promise and reality – How to ensure that the service providers won't use the data for other than what the users are given permission to
- Anonymity technology – How to preserve the anonymity of users throughout the data lifecycle
- Security – How to protect the data throughout its data lifecycle

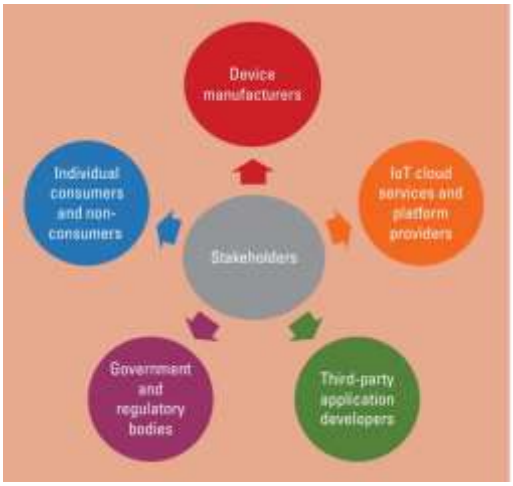

Figure 1. Major stakeholders responsible for protecting user privacy(PERERA ET AL., N.D.)

Device manufacturers must embed some privacy preserving techniques into their own devices. Also must ensure that they implement secure storage, deletion of data and access control. IoT cloud services and platform providers must ensure that they practice the common standards that are accepted in the cloud services industry. Third party application developers have a certain responsibility to make their apps secured. Government bodies and regulatory bodies must lead and enforce standardization and proper legal terms. Individual consumers and non-consumers must be also well concerned about the regulations.

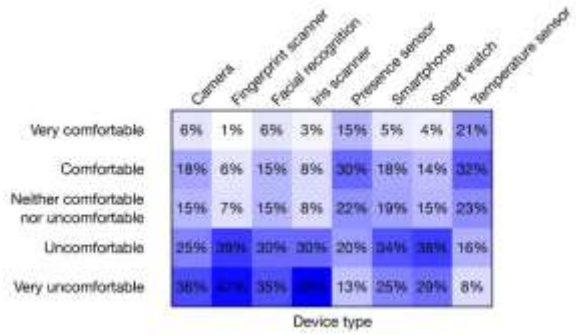

Figure 2. Relation between device type and participants comfort level(EMAMI-NAEINI ET AL., N.D.)

So in this study 1104 participants were analysed with their comfort level and the sensor on which the data are collected. Most of the participants felt very uncomfortable with storing their biometric data using the sensors.(Emami-Naeini et al., n.d.)

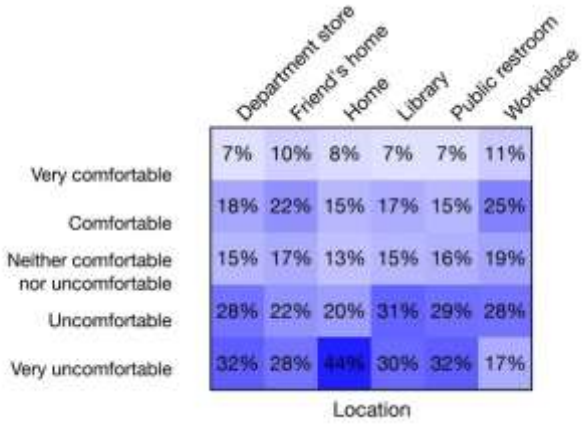

Figure 3. Relation between location data and participants comfort level(EMAMI-NAEINI ET AL., N.D.)

So in this most of the user are felt mostly uncomfortable with storing there home as a location data. Likewise, there are many statistical evidences gathered by researches proving some certain privacy concerns of real life situations tied with users. These concerns will be evaluated in developing the conclusion of this study.

B. Privacy in mobile applications

App publishers must provide a privacy policy and notify users about there app's privacy practices. But the users are totally uncomfortable with the sheer number of permissions associated with most apps. Privacy policies in the smartphone apps are intended to describe the app's

data collection and usage practices. So some researchers had carried out researches like out of the total apps in the playstore how many of those are comprise of a privacy policy. By using crawling mechanisms they had identified the apps that consist of a privacy policy(Lin et al., 2012). Only 63.1% of the apps on the playstore are link to a privacy policy. But they had crawled the playstore in three successive periods with a gap of two to three months and they had noticed the google's actions are contributing to an increase in the percent of apps with a privacy policy.

Based on (Liu et al., 2014) 4.8 million users the research was carried out. And each user had an average of 22 apps installed. Some users agreed on certain permissions and some have a diverging preference.

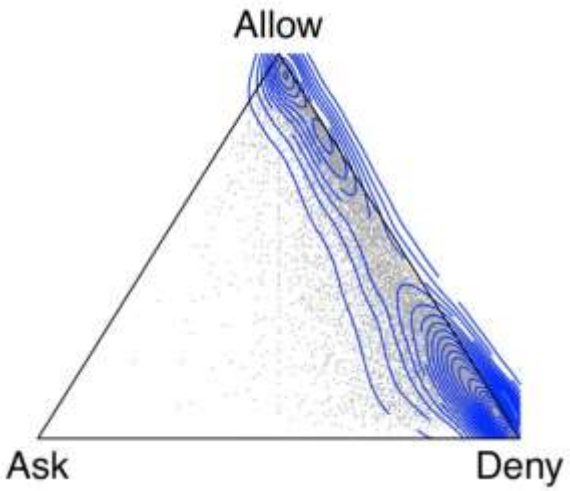

Figure 4. Distribution of user's decision(Allow, Deny, Ask) for each app permission(LIU ET AL., 2014)

In the figure the top corner corresponds to 100% of users 'allow' an app permission and bottom left correspond to 100% of users 'ask' to be prompted for an app permission and bottom right correspond to users select 'deny'. While many dots, are concentrated on top and bottom right corners. So the users have a more bias towards either granting permission or denying the permission. This shows that simple classifiers could be built to predict the users app permission decisions.

Another research (Zimmeck et al., 2017) was carried out to analyse and predict the compliance with the privacy requirements. There analysis was carried out with the base of 17991 free android apps. It was carried out with the combination of machine learning based privacy policy analysis with the static code analysis of apps. On there results 71% lack a privacy policy and 9050 apps that has a certain policy also had many inconsistencies between the app policy and the internal functioning determined through the static code analysis.

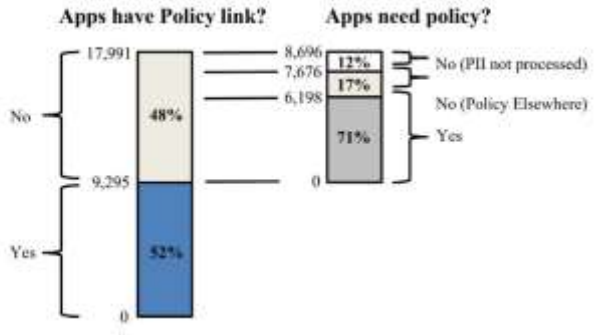

Figure 5. Analysis results of the total apps with the no. of apps having a privacy policy etc.(ZIMMECK ET AL., 2017)

It's really hard to verify whether an application behaves exactly according to the law and its privacy policy. Some researches had used some machine learning techniques and static analysis to identify these inconsistencies.

It mainly advances in 3 main areas such as transparency of data practices, more scalability as a whole eco system and automating the mobile app privacy compliance analysis.

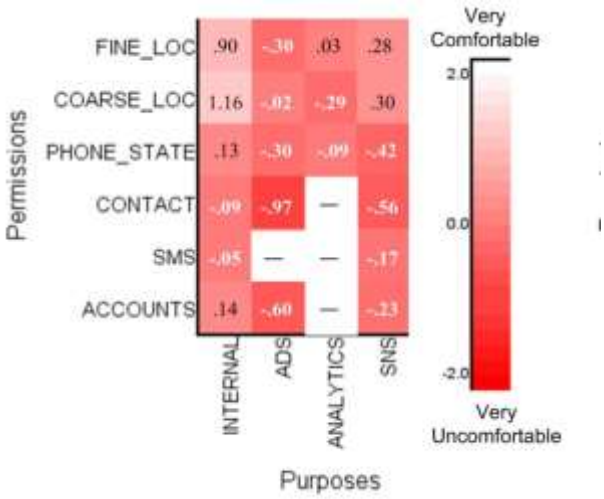

Figure 6. Average self-reported comfort ratings of different permission usages(LIN ET AL., N.D.)

We can obtain a great insight into the users level of concernment in regarding to permissions by going through the Figure 6. Using contact details for the purpose of serving advertisement is the major threat that led the users to be uncomfortable. Users are very less concerned for using for the purpose of analytics.

C. Designing Personalized Privacy Assistant

Android operating systems had introduced new mechanisms to inform users about the data accessed by apps and give them a certain degree of control. Just-in-time privacy nudges are there to inform the users by

using a permission manager running under the hood but this is totally ineffective. Since the users won't have a time to go through all these and tweak the settings according to their preferences(Gibler et al., 2012). So the aim of this study is to propose an effective and intuitive Personalized Privacy Assistant that can actively support the users.

First a field study was conducted by letting a set of android users to use their phones for a period of 1 week and they were provided with privacy nudges daily to increase the awareness and let them tweak the settings over the course of 1 week. Then profiles were build based upon them. After that data was analysed using machine learning algorithms and they were implemented in to the solution Personalized Privacy Assistant. So this is capable of capturing user preferences and suggesting personalized recommendation of app permission settings. Figure 7 depicts the number of recommendations notified to the user and the ones that were accepted. Overall user accept 73.7% of all recommendations and only 5.6% of settings were changed back. Therefore, there is a high satisfactions level among the users towards this Personalized Privacy Assistant.

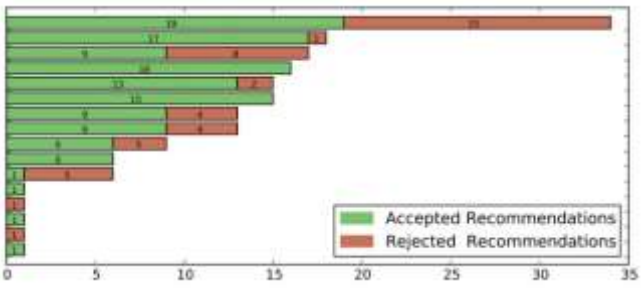

Figure 7. No. of recommendations prompted by the Personalized Privacy Assistant accepted or rejected rate

In this study there are two teams known as control and treatment team. The data are collected regarding the preferences over a week from the control team and then those collected data are analysed using machine learning techniques. The generated model was applied to the Personalized Privacy Assistant and results were tested out with the treatment team.

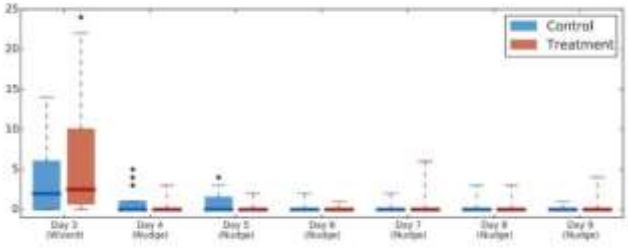

Figure 8. Number of permission changes done by the both groups within the study period

Figure 8 depicts the number of permission changes that the whole team had undergone over the course of 9 days period. Then figure 9 depicts the study flow used in throughout this study

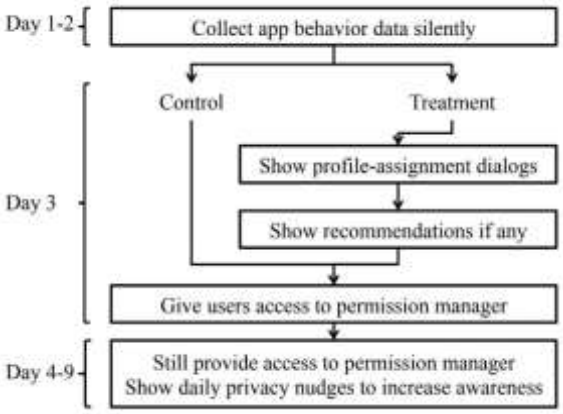

Figure 9. Overview of the study flow in this study(Liu et al., n.d.)

Personalized privacy assistant is envisioned as the best solution for this. This intelligent assistant will be able to learn the privacy preferences of the users over time and with the time it will gain the ability to semi automatically configure many privacy related preferences and settings. The assistant will also be capable of alerting the user about certain preferences or settings that the user may not feel comfortable base on the concernment level model generated as an overview of user preferences over the time. Privacy assistant will help their users to give better insight into the process behind the data processing and empower them to control such processing in the best possible way. So in the end the Personalized Privacy Assistant will be more intuitive and effective manager of privacy preferences with limited number of interruptions and guidance.

IV. CONCLUSION

By analysing the final results of the above studies and the research carried out regarding to these rising privacy issues we can conclude that Personalized Privacy Assistants is the optimum solution for all these concerns. Based on the studies this paper has summarized the most important features that the Privacy Assistant must include for the android platform.

- Organize the permission settings into set of universal user profiles
- Learn the privacy preferences of the users over the time
- Based on the learning ability, suggest the best privacy nudges to determine the concernment level
- Determining the unique concernment level metric of the users
- Gaining the ability to semi-automatically configure the privacy preferences based on the concernment level
- Reduce the lack of awareness among the users regarding the privacy concerns

- Studying the interaction patterns of the user and refine the concernment level
- Ensure the collected data of the user's interactions and preferences are stored securely
- Presenting an insight of the data scraping level of applications and services

With the evolvement of smartphones, Internet of things and online tracking there are many privacy issues that arise. So the society is in need of an intuitive, effective, easy to manage solution to alleviate these issues with the minimum effort. The best solution is an introduction of a fully fledged Personalized Privacy Assistant. The assistant should comprise of the features that are listed above in the evaluation section.

The work can be further improved by studying, analysing the newly emerging privacy concerns and issues.
By refining the machine learning algorithms and models the features can be improved further to make the solution more effective.

REFERENCES

Atzori, L., Iera, A., Morabito, G., 2010. The Internet of Things: A survey. Computer Networks 54, 2787–2805. https://doi.org/10.1016/j.comnet.2010.05.010

Balebako, R., Marsh, A., Lin, J., Hong, J., Faith Cranor, L., 2014. The Privacy and Security Behaviors of Smartphone App Developers, in: Proceedings 2014 Workshop on Usable Security. Presented at the Workshop on Usable Security, Internet Society, San Diego, CA. https://doi.org/10.14722/usec.2014.23006

Emami-Naeini, P., Bhagavatula, S., Habib, H., Degeling, M., Bauer, L., Cranor, L.F., Sadeh, N., n.d. Privacy Expectations and Preferences in an IoT World 15.

Gibler, C., Crussell, J., Erickson, J., Chen, H., 2012. AndroidLeaks: Automatically Detecting Potential Privacy Leaks in Android Applications on a Large Scale, in: Katzenbeisser, S., Weippl, E., Camp, L.J., Volkamer, M., Reiter, M., Zhang, X. (Eds.), Trust and Trustworthy Computing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 291–307. https://doi.org/10.1007/978-3-642-30921-2_17

Joinson, A., Reips, U.-D., Buchanan, T., Schofield, C.B.P., 2010. Privacy, Trust, and Self-Disclosure Online. Human-Computer Interaction 25, 1–24. https://doi.org/10.1080/07370020903586662

Leon, P.G., Ur, B., Wang, Y., Sleeper, M., Balebako, R., Shay, R., Bauer, L., Christodorescu, M., Cranor, L.F., 2013. What matters to users?: factors that affect users' willingness to share information with online advertisers, in: Proceedings of the Ninth Symposium on Usable Privacy and Security - SOUPS '13. Presented at the the Ninth Symposium, ACM Press, Newcastle, United Kingdom, p. 1. https://doi.org/10.1145/2501604.2501611

Lin, J., Liu, B., Sadeh, N., Hong, J.I., n.d. Modeling Users' Mobile App Privacy Preferences: Restoring Usability in a Sea of Permission Settings 14.

Lin, J., Sadeh, N., Amini, S., Lindqvist, J., Hong, J.I., Zhang, J., 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing, in: Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12. Presented at the the 2012 ACM Conference, ACM Press, Pittsburgh, Pennsylvania, p. 501. https://doi.org/10.1145/2370216.2370290

Liu, B., Andersen, M.S., Schaub, F., Almuhimedi, H., Zhang, S.,Sadeh, N., Acquisti, A., Agarwal, Y., n.d. Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions 16.

Liu, B., Lin, J., Sadeh, N., 2014. Reconciling mobile app privacy and usability on smartphones: could user privacy profiles help?, in: Proceedings of the 23rd International Conference on World Wide Web - WWW '14. Presented at the the 23rd international conference, ACM Press, Seoul, Korea, pp. 201–212. https://doi.org/10.1145/2566486.2568035

Melicher, W., Sharif, M., Tan, J., Bauer, L., Christodorescu, M., Leon, P.G., 2016. (Do Not) Track Me Sometimes: Users' Contextual Preferences for Web Tracking. Proceedings on Privacy Enhancing Technologies 2016, 135–154. https://doi.org/10.1515/popets-2016-0009

Perera, C., Wang, L., Zomaya, A.Y., n.d. Privacy of Big Data in the Internet of Things Era 13.

Zimmeck, S., Wang, Z., Zou, L., Iyengar, R., Liu, B., Schaub, F., Wilson, S., Sadeh, N., Bellovin, S.M., Reidenberg, J., 2017. Automated Analysis of Privacy Requirements for Mobile Apps, in: Proceedings 2017 Network and Distributed System Security Symposium. Presented at the Network and Distributed System Security Symposium, Internet Society, San Diego, CA. https://doi.org/10.14722/ndss.2017.23034

ACKNOWLEDGMENT

I would like to express my great appreciation to Mr. PPNV Kumara for his valuable and constructive suggestions during the planning and development of this research work. I would also like to thank Mr. LP Kalansooriya and staff of the Department of Computer Science at KDU, for their valuable and precious time, which is generously and highly admired.